any two steps in the method of claim 1. Claims 2-7 have been amended to clarify that the residues or sequences referred to relate to amino acids.

New claims 17 - 21 have been added. These claims are alternate approaches to claiming the subject matter claimed in claim 1. Additional support for claims 20-21 is found in Example 1.

It is believed that these amendments introduce no new matter. The inventors respectfully request entry of these amendments.

The rejection under 35 U.S.C. § 112, second paragraph

Claims 1-10 were rejected under 35 U.S.C. § 112, second paragraph for indefiniteness. Claims 1, 3, 4, 5 and 7 were rejected for use of the term "import antibody". These claims have been amended to indicate that the import antibody is a non-human antibody which is desired to be humanized.

Claim 1 step (a) was rejected because of the term "consensus human variable domain". The terms "consensus sequence", "consensus antibody" and "consensus human variable domain" are defined at specification page 16, line 29 to page 17, line 17:

> The terms "consensus sequence" and "consensus antibody" as used herein refers to an amino acid sequence which comprises the most frequently occurring amino acid residues at each location in all human immunoglobulins of any particular subclass. In preferred embodiments, the consensus human variable domain sequences are derived from the most abundant subclasses in the sequence compilation of Kabat *et al., Sequences of Proteins of Immunological Interest*, National Institutes of Health, Bethesda MD (1987), namely $V_L$ κ subgroup I and $V_H$ group III....

As described in the specification, a "consensus human variable domain" would have an amino acid sequence comprising, amino acid residue by residue, the most frequently occurring amino acid residue gathered from a group of human immunoglobulins. The identity of each amino acid residue making up the consensus sequence is determined separately, requiring merely routine tabulation of the amino acids present in each member of a particular immunoglobulin subclass. To expedite the routine tabulation of the most commonly occurring amino acids, workers in the field are referred to the Kabat *et al.* publication cited in the quoted material above, which presents such tabulations.

Claim 1 step (d) was rejected as indefinite as to whether the alignment of the amino acid sequences is a physical or mental step. This rejection is somewhat confusing. The inventors intend claim 1, step (d) to refer to a maximal homology alignment of representations of amino acid sequences, as described in the specification at page 17, lines 18-27. Preparing such a homology alignment typically combines physical and mental actions. This connotation for the phrase "alignment of sequences" is common in the art to which this invention pertains. Step (d) of claim 1 does not require

manipulation of the actual, tangible amino acids, merely manipulation of symbolic representations of the actual amino acids.

Claim 1 step (e) was rejected because of the term "homology"; the Examiner questioned whether conservative amino acids are to be considered as homologs. Identity or homology with respect to a specified amino acid sequence of this invention is defined on page 17, lines 18-27. At lines 22-23, the specification indicates that this invention does "not consider[ing] any conservative substitutions as part of the sequence identity". Conservative substitutions are therefore not considered as homologs.

Claim 1 step (f) was rejected for use of the language "participates". Step (f) of claim 1 refers to an amino acid residue which "participates in the in the $V_L$ - $V_H$ interface". This step has been amended to clarify that immunoglobulin residues which so participate are those that affect the proximity or orientation of the $V_L$ and $V_H$ regions with respect to one another.

Claim 1 step (f) was also rejected as indefinite as to how one of ordinary skill can determine the effects listed in steps 1-3. Steps 1-3 presently list the following effects an import amino acid residue might have:

1. non-covalently binds antigen directly,
2. interacts with a CDR; or
3. participates in the $V_L$ - $V_H$ interface by affecting the proximity or orientation of the $V_L$ and $V_H$ regions with respect to one another.

The specification discusses, at pages 13-16, the interactions of amino acid residues within an immunoglobulin and describes at least two methods for evaluating the role of any particular amino acid residue: three dimensional models and assays. As stated at page 14, lines 2-9:

> "Three dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate immunoglobulin sequence, i.e., the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen."

The specification provides detailed information how to evaluate the three-dimensional models to determine the various potential effects of amino acid residue changes.

The specification also suggests an alternate method for evaluating the effect of an amino acid residue change. On page 16, lines 14-18, the specification teaches:

> "Since it is not entirely possible to predict in advance what the exact impact of a given substitution will be it may be necessary to make the substitution and assay the candidate antibody for the desired characteristic. These steps, however, are *per se*

routine and well within the ordinary skill of the art."

The inventors submit that methods for determining the effects of amino acid changes are known in the art, and that those skilled in the art would understand what is claimed in step (f).

Claim 1 step (g) was rejected as being indefinite as to what effects are reasonably expected to occur. The word "reasonably" has been deleted from the claim.

Claim 2 was rejected as lacking antecedent basis for "the domain". This claim has been amended to clarify that the intended domain is the consensus human antibody variable domain.

Claims 3 and 4 were rejected as indefinite for not specifying when in the process one would search for the glycosylation sites. These claims have been amended to indicate that one would search for glycosylation sites between any two steps in the method of claim 1.

Claim 5 was rejected as unclear in the use of the phrase "preparing a humanized antibody"; this phrase has now been added by amendment as the last step of claim 1. The phrase is intended to mean the physical making of a humanized antibody, methods for which are described in the specification, including *in vitro* mutagenesis and recombinant engineering. The Examiner also seems to be questioning how claim 5 differs from the previous claims. Claim 5 adds an additional step of determining if a particular amino acid residue in the consensus human variable domain--which differs from the import antibody amino acid residue at that site--also appears at that site in antibodies of other species at that particular site (is conserved). If the particular amino acid residue is conserved across species at that site, than that residue is retained in the humanized antibody, and not substituted by the import antibody amino acid residue at that site, and without requiring evaluation of the impact of such a change on the antibody's characteristics.

Claim 6 was rejected as vague for unclear use of numbers. These numbers refer to particular amino acids in the light (L) and heavy (H) chains of immunoglobulins. By convention, workers in this field generally utilize the immunoglobulin numbering system set forth in Kabat, E. A. *et al.*, *Sequences of Proteins of Immunological Interest* (National Institutes of Health, Bethesda, MD, 1987)), as described in the specification at page 8, lines 19-21. The Examiner's attention is drawn to Queen *et al.*, already of record in this case, especially at page 10032 column 1 first paragraph (and reference 38 therein) where antibody amino acid residues are referred to with numbers representing certain positions. It is submitted that workers in the field will understand clearly what is claimed in claim 5.

Claim 7 was rejected as indefinite as to what the method is drawn, and has been amended according to the Examiner's suggestion.

According to the CAFC, a decision as to whether a claim is invalid for indefiniteness "requires a determination whether those skilled in the art would understand what is claimed", Amgen v. Chugai,

18 USPQ2d 1116, 1030 (CAFC 1991). The presently pending claims use terminology with clear meanings in the field, especially in light of the definitions provided in the specification. The wordings of the claims comply with the requirements of 35 USC § 112, and this rejection should be reconsidered and withdrawn.

The rejection under 35 U.S.C. § 112, first paragraph

Claims 1-11 were rejected under 35 U.S.C. § 112, first paragraph as lacking enablement.

Claims 1 and 7 were rejected as lacking enablement in the language "at least a portion of an import variable domain". These terms have been deleted from the claims.

Claim 1 step (c) was rejected for being unclear as to how one would determine which amino acids are to be substituted. This step recites "substituting an import CDR amino acid sequence for the corresponding human CDR amino acid sequence". The identification of the Complementarity Determining Region (CDR) amino acid sequence of the import and the human amino variable domain sequences is made in previous step (b). To accomplish step (c), therefore, one substitutes the amino acids identified in step (b).

Methods for identifying CDRs and distinguishing them from Framework Residues (FRs) are known in the art. As the specification describes on page two, antibody variable domains of natural light and heavy chains have the same general structure, and each domain comprises four framework (FR) regions, whose sequences are somewhat conserved, connected by three hyper-variable or complementarity determining regions (CDRs) (see Kabat, E. A. *et al.*, *Sequences of Proteins of Immunological Interest*, National Institutes of Health, Bethesda, MD, (1987)). The four framework regions largely adopt a $\beta$-sheet conformation and the CDRs form loops connecting, and in some cases forming part of, the $\beta$-sheet structure. The CDRs in each chain are held in close proximity by the framework regions and, with the CDRs from the other chain, contribute to the formation of the antigen binding site. The CDR may be identified following three-dimensional modeling of the antibody. The CDRs may also be identified based on comparison of the an antibody amino acid sequence with a known antibody.

Attached as Exhibit A for the Examiner's convenience are pages from the Introduction to Kabat, E. A. *et al.*, *Sequences of Proteins of Immunological Interest, Fifth Edition*, National Institutes of Health, Bethesda, MD, (1991). This work, along with the earlier Kabat compendiums referred to in the specification and other references, guide the practitioner in the numbering of antibody amino acid sequences, and the assignment of particular amino acids to one of the FR or CDR regions. The Examiner's attention is drawn the sections beginning on page xv, the section entitled "Variable Region Sequence" and especially to Table I, page xvi. Table I presents the amino acid residues associated

with FRs and CDRs of the variable domains of immunoglobulin light and heavy chains. See also Figure 1. page xviii, which shows a schematic view of an immunoglobulin; please note the mention in that figure description to the use of a maximum homology alignment to determine the proper numbering of the amino acids (as referred to in the response to the previous § 112 rejection). The inventors submit that the identification of immunoglobulin amino acid residues as belonging to a CDR or to the framework is routine in the art, requiring no undue experimentation.

The specification teaches, in detail, several ways to substitute amino acid residues, including mutagenesis and the construction of nucleic acid encoding the desired sequence. Alanine scanning mutagenesis is described at page 36, line 20 to page 37, line 3. Oligonucleotide-mediated mutagenesis, PCT mutagenesis and cassette mutagenesis are described in the specification at page 39, line 10 through page 44, line 10. The inventors submit that steps (b) and (c) of claim 1 are fully enabled by the specification.

Claim 1 step (f) was rejected as lacking enablement for determining which amino acid residues may be expected to interact with the antigen. At page 29, lines 4-10, the specification teaches that:

> "Differences between the non-human import and the human consensus framework residues are individually investigated to determine their possible influence on CDR conformation and/or binding to antigen. Investigation of such possible influences is desirably performed through modeling, by examination of the characteristics of the amino acids at particular locations, or determined experimentally through evaluating the effects of substitution or mutagenesis of particular amino acids."

Techniques for molecular modeling are described on pages 27-28. Experimental evaluation of the role of particular amino acids will utilize assays tailored to the activities of the antibody to be humanized.

More detailed teaching on identifying residues that influence antigen binding is contained in the specification at page 14, line 10 through page 15, line 6, where it is stated:

> "A residue that noncovalently directly binds to antigen is one that, by three dimensional analysis, is reasonably expected to noncovalently directly bind to antigen. Typically, it is necessary to impute the position of antigen from the spatial location of neighboring CDRs and the dimensions and structure of the target antigen. In general, only those humanized antibody residues that are capable of forming salt bridges, hydrogen bonds, or hydrophobic interactions are likely to be involved in non-covalent antigen binding, however residues which are separated spatially by 3.2 Angstroms or less may also non-covalently interact. Such residues typically are the relatively larger amino acids, such as tyrosine, arginine, and lysine. Antigen-binding FR residues also typically will

have side chains that are oriented into an envelope surrounding the solvent oriented face of a CDR which extends about 7 Angstroms into the solvent from the CDR domain and about 7 Angstroms on either side of the CDR domain, again as visualized by three dimensional modeling.

The inventors submit that determining whether a residue may be expected to influence antigen binding is routine in the art, in light of the detailed teachings of the specification.

Claim 2 was rejected as lacking enablement for determining which residues are exposed on the surface or buried within the domain. As indicated in the specification, for example at page 91, lines 18-21, the worker in this field would examine the structural models of the import and human sequences to determine if an amino acid residue is exposed on the surface of the domain or is buried within. Evaluation of structural models, preparation of which are described in the specification, to determine whether a residue is exposed or buried is routine and within the ordinary skill in the art.

Claim 3 was rejected as lacking enablement for how one would determine which glycosylation site affects antigen binding, or what comprises "reasonable expectation". The specification teaches, at page 8, lines 22-32, teaches that determining if the glycosylation is reasonably expected to be important for the desired antigen binding and biological activity of the antibody involves determining if the glycosylation site binds to antigen or changes a side chain of an amino acid residue that binds to antigen, or if the glycosylation enhances or weakens antigen binding, or is important for maintaining antibody affinity. As with other aspects of this invention, evaluation of the impact of glycosylation typically is performed by evaluation of molecular models, or experimental evaluation of a modified polypeptide. Such evaluation is routine within the field.

Claims 6, 7 and 9 were rejected as being enabled only with respect to IgG and not other antibody isotypes. The specification, at page 13 lines 14-22, states:

"The humanized antibody will be selected from any class of immunoglobulins, including IgM, IgG, IgD, IgA and IgE, and any isotype, including IgG1, IgG2, IgG3 and IgG4. Usually the constant domain is a complement fixing constant domain where it is desired that the humanized antibody exhibit cytotoxic activity, and the class is typically $IgG_1$. Where such cytotoxic activity is not desirable, the constant domain may be of the $IgG_2$ class. The humanized antibody may comprise sequences from more than one class or isotype, and selecting particular constant domains to optimize desired effector functions is within the ordinary skill in the art."

The Examples presented in the specification involve the use of a $IgG_1$ constant domain. As noted in the specification, specific method steps and illustrative reagents for the use of $IgG_1$ are taught, as well

as their applicability to other antibody isotypes. The inventors respectfully reminds the Examiner that working examples are not a required component of a patent application. As stated in MPEP § 608.01(h), "There is no statutory requirement for the disclosure of a specific example." Thus, the absence of a working example describing particular embodiments of the invention cannot negative the patentability of the invention. The examples included in the specification, which illustrate the preparation of $IgG_1$ antibodies, are representative of the manner in which the invention may be practiced. From reading these examples and the detailed description of the invention, the ordinarily skilled artisan would immediately deduce the applicability of the methods described in the specification to other immunoglobulin isotypes.

The Examiner has not made a prima facie case for the § 112, first paragraph rejections, supplying no basis for her skepticism about the scope of the claims. The burden is on the Examiner to provide evidence to support rejections of this sort. "Mere broad generalizations and allegations are insufficient for holding of non-enablement," Ex parte Goeddel, 5 U.S.P.Q. 1449, 1450 (TTAB 1987).

If the Examiner is only prepared to allow claims to exemplified embodiments, what incentive exists for an inventor to disclose the invention to the public? Trade secret protection obviously would be superior to the following circumstances foreseen by the CCPA in In re Goffe, 191 USPQ 429, 431 (CCPA 1976):

> For all practical purposes, the board would limit appellant to claims involving the specific materials disclosed in the examples, so that a competitor seeking to avoid [literally] infringing the claims would merely have to follow the disclosure in the subsequently-issued patent to find a substitute. However, to provide effective incentives, claims must adequately protect inventors. To demand that the first to disclose shall limit his claims to what he has found will work or to materials which meet the guidelines specified for 'preferred' materials in a process such as the one herein involved would not serve the constitutional purpose of promoting progress in the useful arts. See In re Fuetterer, 50 CCPA 1453, 1462, 319 F.2d 259, 265, 138 USPQ 217, 223 (1963).

For a similar case, see In re Strahilevitz, 212 USPQ 561 (P.O.B.A. 1982), where the Board was reversed for rejecting as non-enabling an application that was devoid of even a single working example.

The first paragraph of 35 U.S.C. § 112 requires nothing more than objective enablement. Whether this is achieved by the use of illustrative examples or by broad terminology is of no importance, In re Marzocchi et al., 169 USPQ 267 (CCPA 1971). Further, an assertion by the Examiner that the enabling disclosure is not commensurate with the protection being sought must be supported by reasons for doubting the truth or accuracy of any statement in the presumptively accurate supporting disclosure. It is also incumbent upon the Examiner to back up such assertions with acceptable evidence or reasoning to substantiate the doubts so expressed, In re Armbruster, 185

USPQ 152 (CCPA 1975), In re Strahilevitz, *op cit*.

Se also In re Smith, *supra*, wherein the CCPA reversed an Office ruling that the description in the specification of two categories of prepolymers was not sufficient to support the broad claim for all polymers having a certain desired property. In this case, the court even acknowledged that the specification did not contain language that was precisely identical to the language of the claims. However, the tenor of the specification was that the applicant had made a generic invention rather than one limited to two categories of polymers.

In the present situation, the Examiner has provided no evidence to support the assertion that the invention is not enabled for the preparation of humanized antibodies. Broad claims should be allowed if there is adequate disclosure and where, as in the present situation, there is no pertinent art to prevent such claims. As stated in In re Sus and Schaefer, 134 USPQ 301, 304 (CCPA 1962) (emphasis added):

> The public purpose on which the patent law rests required the granting of claims commensurate in scope with the invention disclosed. This requires as much the granting of broad claims on broad inventions as it does the granting of more specific claims on more specific inventions. It is neither contemplated by the public purpose of the patent laws nor required by the statute that an inventor shall be forced to accept claims narrower than his invention in order to secure allowance of his patent.

The inventors submit that in view of the detailed information provided in the specification as discussed above, the specification adequately teaches how to practice the claimed invention. The rejections under 35 USC § 112, first paragraph, should be reconsidered and withdrawn, as they are not statutorily based, are inconsistent with court and Patent Office decisions on the subject, and are contrary to public policy.

The rejection under 35 U.S.C. § 101

Claims 1-4, and 6-8 were rejected under 35 U.S.C. § 101 as being directed to non-statutory subject matter. It is believed that the amendments to the claims made above render moot this ground of rejection.

The rejection under 35 U.S.C. § 112, first paragraph and under 35 U.S.C. § 101

Claims 9-13 were rejected under 35 U.S.C. § 112, first paragraph and under 35 U.S.C. § 101 as lacking utility for the treatment of malignant and autoimmune diseases in humans. The inventors request clarification of this rejection, because none of claims 9-13 are directed to methods of treatment. These claims are directed to humanized antibody variable domains and the polypeptides

of recited sequences. These polypeptides are useful as probes, and in diagnostic assays, as described in the specification at pages 65-66, and need not rely on therapeutic utility.

The rejections under 35 U.S.C. § 102(b) and § 102(a)

Claims 1, 2, and 5-10 were rejected under 35 U.S.C. § 102(b) as being anticipated by Queen *et al.*, and under 35 U.S.C. § 102(a) as being anticipated by Co *et al.*. The inventors respectfully traverse these rejections.

To constitute anticipation, all material elements of a claim must be found in one prior art source. In re Marshall, 198 USPQ 344 (CCPA 1978); In re Kalm, 154 USPQ 10 (CCPA 1967). The inventors will show that neither Queen nor Coe contains all the material elements of these claims, particularly the limitation regarding the use of a consensus sequence.

The rejected claims are directed to the humanization an antibody, namely the combination of amino acid sequence from a non-human antibody desired to be humanized, and from a consensus human variable domain. Methods for preparing such a consensus sequence are fully described in the specification and are discussed above. The inventors believe that the use of a such a consensus sequence achieve a superior result, or a "better" humanized antibody.

The cited prior art utilizes a different approach, which approach had apparently been taken by all other workers in the field prior to the present invention. These workers did not prepare a consensus human antibody to combine with their non-human antibody. Instead, they selected only one human antibody for use, based on the similarity of that human antibody to their non-human antibody. Queen *et al.* state this objective explicitly, at page 10031, column 2 of their paper:

> "In selecting a human antibody to provide the variable region framework for the humanized anti-Tac antibody, we reasoned that the more homologous the human antibody was to the original anti-Tac antibody, the less likely would combining the anti-Tac CDRs with the human framework be to introduce distortions into the CDRs."

Queen continues to describe selecting a human heavy chain V region which was 57% identical to their non-human antibody, after dismissing all other candidate as between 30-52% identical to their non-human. They selected the human light chain V region from the same human antibody for their use.

Co *et al.* are equally explicit describing their similar reasoning. At page 2871, column 1 they state:

> "First, a human antibody variable region with maximal homology to the mouse antibody is selected to provide the framework sequence for humanization of the mouse antibody. Normally the heavy chain and light chain from the same human antibody are

chosen so as to reduce the possibility of incompatibility in the assembly of the two chains. Based on a sequence homology search against the NBRF protein sequence data base, the antibody Pom was chosen to provide the framework sequences for humanization of Fd79."

The approach of the present invention is quite distinct, in its use of a consensus human variable domain sequence. This consensus sequence might or might not have a high degree of homology with the non-human antibody. Neither Queen *et al.* or Coe *et al.* supply this teaching, and therefore do not anticipate the claimed invention. This rejection should be reconsidered and withdrawn.

The rejection under 35 U.S.C. § 103

Claims 3 and 4 were rejected under 35 U.S.C. § 103 as being obvious over Queen *et al.* or Co *et al.* in view of Wallick *et al.*. Claim 11 was rejected under 35 U.S.C. § 103 as being unpatentable over Queen *et al.* or Co *et al.* in view of Reichmann *et al.*

None of the cited references teaches or suggests the claimed invention, which involves the preparation humanized antibodies using a consensus human antibody variable domain. Such a method is not suggested in any of the prior references, and absent such a teaching there was no motivation to try the methods described in the present specification.

The Obviousness Rejections Do Not Meet the Test of Graham v. Deere

The proper context for determining the issue of obviousness is provided in the seminal decision of Graham v. John Deere, 383 U.S. 1, 148 U.S.P.Q. 459 (1966). In that case, the U.S. Supreme Court set forth the following considerations for deciding this issue:

(1) The scope and the content of the prior art;

(2) The difference between the prior art and the claims at issue;

(3) The level of ordinary skill in the pertinent art; and

(4) Secondary considerations such as commercial success, long-felt and unresolved needs, failure of others, etc.

a. **Scope and Content of the Prior Art.**

1. Queen *et al.* teach the humanization of an anti-Tac antibody. They do not teach the use of a human consensus variable domain to provide the framework for their non-human CDRs.

2. Co *et al.* teach the humanization of an anti-HSV antibody. They do not teach the use of a human consensus variable domain to provide the framework for their non-human CDRs.

3. Wallick *et al.* teach the importance of glycosylation for maintaining the affinity of a monoclonal antibody for its antigen. They do not teach methods for humanization of antibodies, nor teach the creation of a human antibody variable domain consensus sequence.

4. Reichmann *et al.* teach the humanization of an anti-CAMPATH-1 antibody. They do not teach the creation of a human antibody variable domain consensus sequence, or suggest that such might be desirable to provide the framework for their non-human antibody CDRs.

**b. The Differences Between the Prior Art and the Claims at Issue**

The Examiner has chosen various pieces of prior art and concludes that the combination of these references would have rendered the invention obvious.

The prior art shows that it was known as of the filing date to produce antibody fragments comprising sequence from a non-human antibody and from a human antibody. Prior to the present filing date, however, methods were not known which included the use of a consensus human variable domain for mounting the non-human CDRs. There would have been no impetus on the part of the skilled artisan at the filing date to attempt to produce such a consensus sequence or use it in antibody humanization, in view of the teachings of the prior art literature. The cited references do not teach or suggest the claimed invention, alone or in any combination, nor would there have been any reason from these references to practice the claimed methods. The absence of a suggestion of the claimed invention in the art of record precludes the Patent Office from satisfying its initial burden of showing prima facie obviousness.

c. Level of Ordinary Skill in the Art.

The Graham inquiries point to a conclusion of non-obviousness of the present claims regardless of the presumed level of skill in the art. However, absent evidence to the contrary, a person of ordinary skill in the art is presumed to be one who essentially follows conventional wisdom and does not undertake to innovate. As stated by the Federal Circuit in Standard Co. v. American Cyanamid Co., 227 U.S.P.Q. 293, 298 (Fed. Cir. 1985):

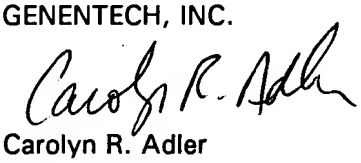
> A person of ordinary skill in the art is also presumed to be one who thinks along the line of convention wisdom in the art and is not one who undertakes to innovate, whether by patient, and often expensive, systematic research or by extraordinary insights, it makes no difference which.

The inventors submit that one who followed the conventional wisdom would not have

extrapolated from the teachings of the cited references methods for using a consensus human antibody variable domain for humanizing a non-human antibody. Such an extension of the prior art teachings is based entirely upon hindsight analysis of the inventors' methods. The teachings of this invention should not be considered sufficient to support a conclusion of obviousness in this regard.

The inventors submit that in light of the foregoing amendments and remarks the subject matter defined by the pending claims is useful, enabled, and patentable over the references relied upon by the Examiner, which in no way teach or suggest the invention. The inventors believe the claims are now in condition for allowance and earnestly solicit a Notice to that effect. If the Examiner has any questions, she should feel free to contact the undersigned attorney at the telephone number indicated above.

Respectfully Submitted,
GENENTECH, INC.



Carolyn R. Adler
Reg. No. 32,324

29 January 1993

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on 29 January 1993.

Dated: 29 January 1993

Carolyn R. Adler

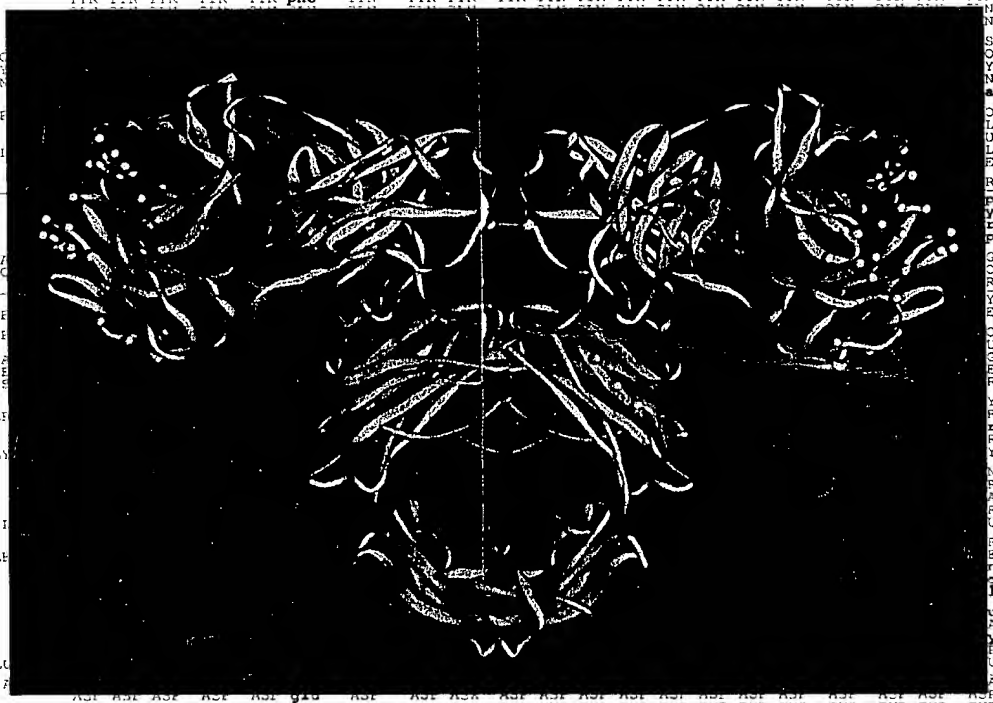
c124.u

Carter

# SEQUENCES OF PROTEINS OF IMMUNOLOGICAL INTEREST

VOL. I

FIFTH EDITION



1991

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
National Institutes of Health

EXHIBIT A

# SEQUENCES OF PROTEINS OF IMMUNOLOGICAL INTEREST

## FIFTH EDITION

Tabulation and Analysis of
Amino Acid and Nucleic Acid Sequences of Precursors,
V-Regions, C-Regions, J-Chain, T-Cell Receptors for Antigen,
T-Cell Surface Antigens, $\beta_2$-Microglobulins,
Major Histocompatibility Antigens, Thy-1, Complement,
C-Reactive Protein, Thymopoietin, Integrins, Post-gamma Globulin,
$\alpha_2$-Macroglobulins, and Other Related Proteins

**1991**

*Elvin A. Kabat*, Tai Te Wu†, Harold M. Perry‡,*
*Kay S. Gottesman‡, and Carl Foeller‡*

*Depts. of Microbiology, Genetics and Development, and Neurology, Cancer Center/Institute of Cancer Research, College of Physicians and Surgeons, Columbia University, New York, NY 10032 and the National Institute of Allergy and Infectious Diseases, and the Office of the Director National Institutes of Health, Bethesda, MD 20892.

†Depts. of Biochemistry, Molecular Biology, and Cell Biology, and Engineering Sciences and Applied Mathematics and Biomedical Engineering, Northwestern University, Evanston, IL 60208 and the Cancer Center, Northwestern University Medical School, Chicago, IL 60611

‡BBN Systems and Technologies, 10 Moulton Street, Cambridge, MA 02138

‡Formerly with BBN. Present address Laboratory for Applied Research in Academic Information, William H. Welch Medical Library, The Johns Hopkins University, Baltimore, Md 21205

The collection and maintenance of this data base is sponsored under grant 5R01 AI-125616 to E.A. Kabat of Columbia University by the following components of the National Institutes of Health, Bethesda, MD 20892:

Office of the Director
National Center for Research Resources
National Cancer Institute
National Institute of Allergy and Infectious Diseases
National Institute of Diabetes, Digestive and Kidney Diseases
National Institute of General Medical Sciences
National Library of Medicine

Work with the PROPHET software package is supported by a subcontract from Columbia University to BBN Systems and Techologies, Cambridge, MA 02138

U.S. DEPARTMENT OF HEALTH
AND HUMAN SERVICES

Public Health Service
National Institutes of Health

NIH Publication No. 91-3242

**INTRODUCTION**

Our earlier "Variable Regions of Immunoglobulin Chains" (1), the second edition "Sequences of Immunoglobulin Chains" (2) and the third edition "Sequences of Proteins of Immunological Interest" (3) have been further exixpanded in the Fourth Edition (4) and now in the Fifth Edition to include amino acid and nucleotide sequences of precursors, variable regions, constant regions, J-chains of immunoglobulins, β2-microglobulins, antigens of the major histocompatibility complex (HLA, H-2, Ia, DR) as well as of Thy-1, complement, T-lymphocyte receptors for antigens, other T-cell antigens of the immunoglobulin superfamily, interleukins, integrins and various other proteins related to immune functions. The identification and sequencing of clones obtained using recombinant DNA techniques has yielded nucleotide sequences of signal, variable, and constant regions of immunoglobulins (5,6), and these nucleotide sequences have been translated into amino acid sequences. The latter have been included in the tables of amino acid sequences with those determined earlier directly by amino acid sequencing and are indicated by an apostrophe followed by CL after the name of the clone. We have continued to use the PROPHET Software Package of the National Center for Research Resources, National Institutes of Health (7,8) to tabulate the sequences.

In compiling the data for this Fifth Edition we have tried to be as up-to-date as possible and have included only sequences which have been published or which have been accepted for publication. Residues which have not been definitely determined have been excluded. It should be remembered that sequences are often published in review articles without detailed documentary evidence. These have often been revised. We have listed such revisions in the notes in many instances; others can readily be found by comparison with sequences in previous editions. We have compiled sequences determined directly as amino acids and have merged with them those translated from the nucleotide sequences thus making all comparable data available. When antibody activities were known, they have been listed after the amino acid and nucleotide sequence tables and are included in the indexes.

When doubts arise as to the validity of any residue in a sequence, the original reference should be examined to ascertain whether definitive evidence for the sequence has been provided. In earlier editions, we have sent the amino acid and nucleotide sequences as stored in the computer to the original authors for verification. If so verified, this was denoted by "checked by author" at the end of each reference and except for the earliest sequences, the date on which the checked sequence was returned to us is given. Whenever possible, nucleotide sequences from GenBank (9) have been used. Programs for converting a GenBank sequence to the codon format of our tables have been developed. The correctness of the table sequence has been verified by converting back into the linear form and comparing with GenBank. When this has been done the sequence is listed as "from GenBank". Recently we have developed newer programs that automatically process a GenBank entry completely - e.g.: extract the relevant feature, determine the appropriate table, and perform alignment. In such cases, the reference will end with "processed automatically from GenBank:" followed by a list of the GenBank accession numbers from which the data was obtained. Some nucleotide sequences were transmitted to us by electronic mail, and they are indicated by "received from authors through e-mail." If the sequences were entered by us from the literature and then checked with GenBank, this is indicated by "checked with GenBank". We have entered many nucleotide sequences which were not then available from GenBank. In general, we have not included stretches of sequence such as enhancers, switch regions and introns. Much information about such sequences may be found in references (10-13). We have also had access to the Protein Information Resource (14) and to the European Molecular Biology Laboratories Data Base (15).

It is also possible, by examining the numbers of sequences at the

end of each table and the summary tables, to evaluate the probability that a given amino acid at a given position may not be correct. This is most readily done for the framework residues of the V-region and for the C-region; in the complementarity-determining regions this is more difficult because of the high variability.

## AMINO ACID SEQUENCES

The first column in each table gives the residue number. Except for complement, T-cell surface antigens, integrins and miscellaneous proteins, the second column is a tabulation of invariant residues. Since exceptions to invariance are found, the frequency, if less than 1.0 and greater than or equal to 0.95, is indicated alongside the residue listed as invariant; when only a single sequence is available, this is not given. These rows are shaded in grey. Each sequence is tabulated in each subsequent column. Three dashes (---) indicate that no amino acid is present at that position and that the sequence continues. In all instances residues considered uncertain by the authors have not been included in the table. In some instances the symbol # is used to indicate that several amino acid residues were found in one position, and these residues are listed in the notes. The four columns at the end of each table give:

1. the number of residues sequenced at that position,
2. the number of different amino acids found at that position,
3. the number of times the most common amino acid occurred and that amino acid in parentheses, and
4. the variability.

These columns are included only in tables with more than five sequences. Miscellaneous tables have only columns corresponding to the first two above.

Variability is calculated (16) as:

$$\text{Variability} = \frac{\text{Number of different amino acids occurring at a given position}}{\text{Frequency of the most common amino acid at that position}}$$

An invariant position would have a variability of one; if 20 amino acids occurred with equal frequency, the variability would be 20 divided by 0.05 equals 400. If, for example, four different amino acids Ser, Asp, Pro, and Thr occurred at a given position, and of 100 sequences available at that position, Ser occurred 80 times, the variability would be 4/0.8 = 5. When any of the amino acid residues, sequenced directly as amino acids, were not identified completely and are listed as Glx (or Asx), two values, separated by a comma, are given in the last three columns. The first value in each of these columns is calculated assuming that only one of the two possibilities, e.g., Glu or Gln (or Asp or Asn) occurred, while the second considers that both were present and maximizes variability. In the variability plots, the horizontal bars indicate the two values.

When two or more amino acids are most common and occur with equal frequency, they are tabulated as a note, and the symbol + is used in the next to last column. If no sequence data have been reported for any position, there are no entries in the last four columns. Variability is not calculated for insertions or if only a single sequence is known. When the translated sequence of a clone corresponds to a previously listed sequence of a plasmacytoma from

which it was prepa
variability computa
If a given sequence
is indicated by an
antibody specificit
constants if avail
rabbit heavy chain
domain of the rabb
sequence is given;
usually the most ne
included, especiall
Notes are of two t
the symbol #, and s

**Signal Sequences**
The signal (precu
chains are listed
light chains, for
total of nine precu
sequencing of signa
sequences from DNA
acid residues in
Genomic DNA clones
the coding sequenc
-4, and in rare cas
leader peptide to
for positions -4 t

The signal amino
antigens, β2-micr
proteins, complem
proteins are liste

By conformational
Leu-Leu-Leu-Trp-Va
alpha helical
conformations in t
four amino termin
(20).

**Variable Region S**
The variable regi
contain hypervari
(27-30) chains,
labeled with hapt
segments of ligh
examination of se
chains aligned f
These and the thre
were hypothesized
regions or segme
contact with vari
high resolution x
been verified by
all antibodies
hypervariable reg
antibody combinin
the framework (
framework segme
complementarity-d
the three CDRs s
Figures 3-47 have
comments are giv
bibliography. The
Table I.

which it was prepared, only one sequence is listed so that the variability computations are not affected, and a note is included. If a given sequence is associated with any antibody activity, this is indicated by an asterisk alongside the protein heading, and the antibody specificities are given in a separate list with binding constants if available. The notes list the a-allotypes for the rabbit heavy chain V-region and the b-allotypes for the constant domain of the rabbit kappa light chain. A key reference to the sequence is given; generally the most recent reference since it is usually the most nearly complete, but often several references are included, especially when revisions of a sequence have been made. Notes are of two types: general notes about a table indicated by the symbol #, and specific notes indicated by the sequence number.

**Signal Sequences**

The signal (precursor) amino acid sequences of immunoglobulin chains are listed as human, mouse, and miscellaneous for kappa light chains, for lambda light chains, and for heavy chains for a total of nine precursor tables. They were obtained either by direct sequencing of signal proteins (17-19) or by translating nucleotide sequences from DNA clones. Signal segments range from 17-29 amino acid residues in length and are thus numbered from -29 to -1. Genomic DNA clones contain introns of varying length that interrupt the coding sequence of the precursor within the codon for position -4, and in rare cases for position -6. Thus, the L-gene encodes the leader peptide to position -4 and the 5' end of the V-gene codes for positions -4 to -1.

The signal amino acid sequences of the T-cell receptors for antigens, β2-microglobulins, major histocompatibility complex proteins, complement components, integrins, and other related proteins are listed in separate tables.

By conformational energy calculations, the core $V_\kappa$ hydrophobic Leu-Leu-Leu-Trp-Val-Leu-Leu-Leu (MOPC321, MOPC63) exists in an alpha helical conformation, terminated by chain reversal conformations in the four C-terminal residues Trp-Val-Pro-Gly; the four amino terminal residues are compatible with the alpha helix (20).

**Variable Region Sequences**

The variable regions (21) of immunoglobulins have been shown to contain hypervariable segments in their light (16,22-26) and heavy (27-30) chains, of which certain residues have been affinity labeled with haptenic determinants (31-44). Three hypervariable segments of light chain were delineated from a statistical examination of sequences of human $V_\kappa$, human $V_\lambda$, and mouse $V_\kappa$ light chains aligned for maximum sequence similarity (16,23,24,27). These and the three corresponding segments of the heavy chains (27) were hypothesized (16,27) to be the complementarity-determining regions or segments (CDR) containing the residues which make contact with various antigenic determinants, several years before high resolution x-ray structures were determined, and this has now been verified by X-ray diffraction studies at high resolution for all antibodies examined Figures 3-47. The proposed fourth hypervariable region (cf. 30) of heavy chains is not part of the antibody combining site (27). The rest of the V-region constitutes the framework (16,27,45-54). It is convenient to identify the framework segments (FR1, FR2, FR3, and FR4) and the complementarity-determining segments (CDR1, CDR2, and CDR3) with the three CDRs separating the four FRs. The CDRs in the stereo Figures 3-47 have solid circles for each residue. References and comments are given with each figure and are not listed in the bibliography. The residue numbers for these segments are given in Table I.

**TABLE I**

Amino Acid Residues Associated with Framework(FR) and Complementarity Determining Regions (CDR) of the Variable Domains of Immunoglobulin Light ($V_L$) and Heavy ($V_H$) Chains

| Segment | Light Chain | Heavy Chain |
|---|---|---|
| FR1 | 1-23 (with an occasional residue at 0, and a deletion at 10 in $V_\lambda$ chains) | 1-30 (with an occasional residue at 0) |
| CDR1 | 24-34 (with possible insertions numbered as 27A,B,C,D,E,F) | 31-35(with possible insertions numbered as 35A,B) |
| FR2[a] | 35-49[a] | 36-49 |
| CDR2 | 50-56 | 50-65 (with possible insertions numbered as 52A,B,C)[b] |
| FR3 | 57-88 | 66-94 (with possible insertions numbered as 82A,B,C) |
| CDR3 | 89-97 (with possible insertions numbered as 95A,B,C,D,E,F) | 95-102 (with possible insertions numbered as 100A,B,C,D,E,F,G,H,I,J,K) |
| FR4 | 98-107 (with a possible insertion numbered as 106A) | 103-113 |

[a] Five Basilea rabbits (λ) immunized with type II pneumococci and which produced anti-type II pneumococcal polysaccharide had Met at position 48 and an insertion of four amino acid residues between positions 48 and 49; in four of the five the sequence was Glu, Leu, Lys, Ser and the fifth was Trp, Leu, Arg, Lys (53,54,63,64); the others were not sequenced at these positions (for references see table of rabbit λ amino acid sequences.)

[b] In the rabbit, Mage et al. (65) consider position 65 in $V_H$ to be in FR3, since it is allotype related.

The V-genes for the
and the J-minigenes f
kappa light chains. F
by recombination and
by the J-minigene.
occur at different po
residues may result a
of the inserted resi
for better alignment
the V-gene region. In
times more frequentl

The V-genes for the h
and are followed b
extensive variation
ability to be read
boundary between D a
acid position. In adc
sequences vary by a f
of D-J joining appea
between V and D and
and correlates with t
B cells (60). The or
has therefore been re
evidence suggesting
perhaps a minigene
nucleotides. Light c
$V_L$-$J_L$ junction (62),
probably results fro
in fetal and neonata
and 17/146 RNA seque
lower than in adults
regulated both in T
diversity but are te

In the tables of V
horizontal lines for
chain, MPC 11, has
between position 1
have internal deleti

The V-genes for the light chains code to amino acid position 95, and the J-minigenes from position 97 to 107 for lambda and 108 for kappa light chains. Position 96 is usually the site of V-J joining by recombination and may be coded partly by the V-gene and partly by the J-minigene. Because the site of V-J recombination could occur at different positions within a codon, different amino acid residues may result at this position. We have changed the location of the inserted residues from 97A-F (2) to 95A-F, since it makes for better alignment by confining chains of different lengths to the V-gene region. In mouse $V_\kappa$ chains, J1 and J2 were used 5 to 10 times more frequently than J4 and J5 (55).

The V-genes for the heavy chains code up to amino acid position 94 and are followed by the D- and J-minigenes. Because of the extensive variation in the lengths of D-minigenes, and their ability to be read in different reading frames (56), the exact boundary between D and J is not always located at the same amino acid position. In addition, the lengths of the J encoded amino acid sequences vary by a few amino acid residues. Moreover, the process of D-J joining appears to involve insertions of extra nucleotides between V and D and between D and J, termed the N region (57-61) and correlates with the appearance of terminal deoxytransferase in B cells (60). The original numbering system for the heavy chains has therefore been retained. Wysocki et al. (61) have provided some evidence suggesting a non-random origin for the $V_H$-$D_H$ junction, perhaps a minigene, rather than random addition of the N nucleotides. Light chains do not appear to have N sequences at the $V_L$-$J_L$ junction (62), but show an additional residue 95A which probably results from $V_L$-$J_L$ joining. N sequences are generally rare in fetal and neonatal mouse $V_H$-D-$J_H$ junctions (62), only 1/87 DNA and 17/146 RNA sequences contained N regions, an incidence much lower than in adults indicating that N insertion is developmentally regulated both in T and B cells. P elements also contribute to diversity but are templated (62a).

In the tables of V-regions, the FR and CDR are separated by horizontal lines for convenience in reading. One mouse kappa light chain, MPC 11, has an extra segment of 12 amino acid residues between position 1 and the signal sequence (66). Several chains have internal deletions.

Figure 1 (50) shows the domain structure for IgG1 protein EU. Numbering on the left half indicates the CDR for the light and heavy chains (50), while that on the right half gives the EU numbering (67).

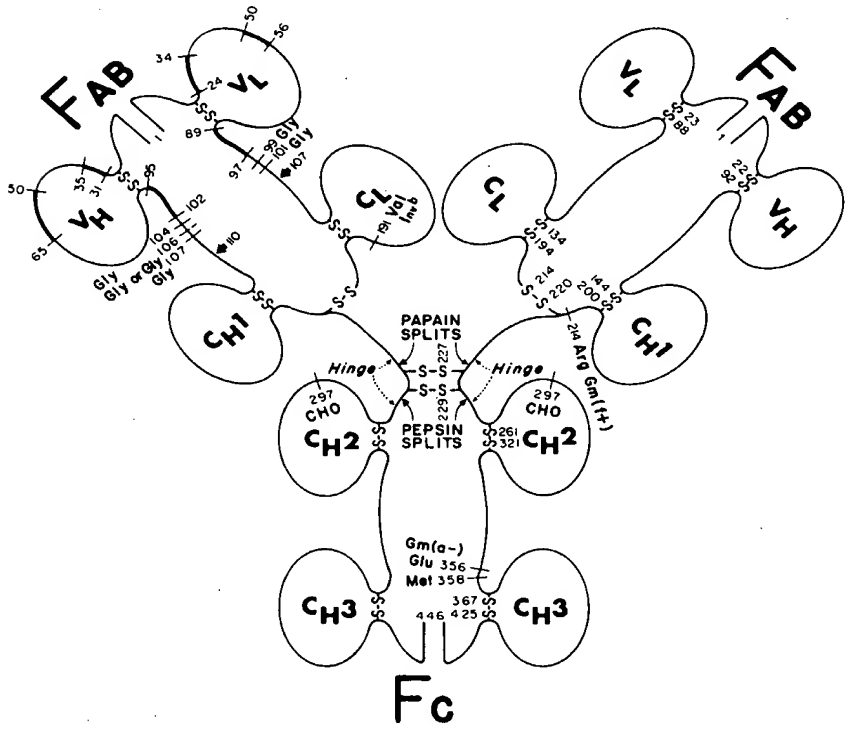



FIG. 1. Schematic view of four-chain structure of human IgG1$_\kappa$ molecule. *Numbers on right side:* actual residue numbers in protein EU [Edelman *et al.* (67)]; *Numbers* of Fab fragment on *left side* aligned for maximum homology; light chains numbered as in Wu and Kabat (16) and heavy chains as in Kabat and Wu (27). Heavy chains of EU have residue 52A, three residues 82A,B,C, and lack residues termed 100A,B,C,D,E,F,G,H,I,J,K, and 35A,B. Thus residue 110 (end of variable region) is 114 in actual sequence. Hypervariable regions, complementarity-determining segments or regions (CDR): *heavier lines.* $V_L$ and $V_H$: light and heavy chain variable region; $C_H1$, $C_H2$, $C_H3$: domains of constant region of heavy chain; $C_L$: constant region of light chain. *Hinge region* in which two heavy chains are linked by disulfide bonds is indicated approximately. Attachment of carbohydrate is at residue 297. *Arrows* at residues 107 and 110 denote transition from variable to constant regions. Sites of action of papain and pepsin and locations of a number of genetic factors are given. Modified from 50.

Critical understanding
sites and the genetic
antibody complementar
evaluation of a large
and especially of the c
and heavy chains of im
to locate residues i
determinants (68,69) a
combining sites will d
and scope $V_H$ and $V_L$ cha
must be resolved. T
immunochemical data in
in addition to other m
resolution X-ray cryst

Through the generous co
been provided with the
Fab molecules, $V_H$ dime
Drs. Eduardo Padlan a
shown. Legends and k
model.

Critical understanding of the architecture of antibody combining sites and the genetics of the generation of diversity and of antibody complementarity depends to a great extent on the evaluation of a large number of sequences of the variable regions and especially of the complementarity-determining segments of light and heavy chains of immunoglobulins of different species. Ability to locate residues in the site making contact with antigenic determinants (68,69) and to predict (70) the structures of antibody combining sites will depend heavily upon such sequences. The role and scope $V_H$ and $V_L$ chains in contributing to binding of the epitope must be resolved. This can be often accomplished by use of immunochemical data in defining antibody combining sites (68,70-73) in addition to other methodologies such as 2D-NMR (71,51) or high resolution X-ray crystallography.

Through the generous cooperation of X-ray crystallographers we have been provided with the α-carbon coordinates of almost all available Fab molecules, $V_H$ dimers and antigen-antibody complexes from which Drs. Eduardo Padlan and Chantal Abergal made the stereo models shown. Legends and key references for each are listed with the model.

E. Long, R. Mage, C. ub and L. Steiner for rlier editions.

s of Fab and antigen as α-carbon coordinates in the text where these an and Chantal Abergel (stereodiagrams from ndebted to them. Dr. the x-ray structure of bted to Drs. Rose Mage the introduction and to tions to organizing the Margulies made helpful text. Dr. Lawrence A. se $V_H$ and $V_\kappa$ families as $V_H$ and $V_\kappa$ sequences into

## REFERENCES TO INTRODUCTION

1. Kabat, E.A., Wu, T.T., and Bilofsky, H. (1976) Variable Regions of Immunoglobulin Chains. Medical Computer Systems, Bolt Beranek and Newman Inc., Cambridge, MA.

2. Kabat, E.A., Wu, T.T., and Bilofsky, H. (1979) Sequences of Immunoglobulin Chains. National Institutes of Health, NIH Publication 80-2008.

3. Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller, M., and Perry, H. (1983) Sequences of Proteins of Immunological Interest. National Institutes of Health.

4. Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H.M., and Gottesman, K.S. Sequences of Proteins of Immunological Interest. (1987) U.S. Dept. of Health and Human Services, Public Health Service National Institutes of Health.

5. Tonegawa, S., Maxam, A.M., Tizard, R., Bernard, O., and Gilbert, W. (1978) Sequence of a mouse germ-line gene for a variable region of an immunoglobulin light chain. Proc. Natl. Acad. Sci. U.S.A. 5:1485-1489.

6. Seidman, J.G., Leder, A., Edgell, M.H., Plosky, F., Tilghman, S.M., Tiemeier, D.C., and Leder, P. (1978) Multiple related immunoglobulin variable region genes identified by cloning and sequence analysis. Proc. Natl. Acad. Sci. U.S.A. 75:3881-3885.

7. Raub, W.F. (1974) The PROPHET system and resource sharing. Federation Proceedings 33:2390-2392.

8. Hollister, C. (1988) A national computational resource for life science research. Nucleic Acids Research 16:1873.

9. Burks, C., Fickett, J.W., Goad, W.B., Kanehisa, M., Lewitter, F.I., Rindone, W.P., Swindell, C.D., Tung, C.-S., and Bilofsky, H. (1985) The GenBank nucleic acid sequence database. CABIOS 1:225-233.

10. Trifonov, E.N. and Brendel, V. (1986) GNOMIC, A Dictionary of Genetic Codes. Balaban Publishers; Rehovot, Philadelphia.

11. Nussinov, R. (1986) Some guidelines for identification of recognition sequences: regulatory sequences frequently contain (T)GTG/CAC(A), TGA/TCA and (T)CTC/GAG(A). Biochimica et Biophysica Acta 866:93-108.

12. Ghosh, D. (1990) A relational data base of transcription factors. Nucleic Acids Res. 18:1749-1756.

13. Wu, T.T., Reidmiller, M., Perry, H.M., and Kabat, E.A. (1984) Long identical repeats in the mouse γ2b switch region and their implications for the mechanism of class switching. EMBO J. 3:2033-2040.

14. Protein Information Resource. National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC 20007.

15. European Molecular Biology Organization (1990) EMBL File Server. See issues of Nucleic Acid Research for recent sequences added to their data bank.

16. Wu, T.T. and Kabat, E.A. (1970) An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light [illegible] their implications for antibody complementarity. J.

17. Milstein, C., Brownlee, G.G., Harrison, T.M., and Mathews, M.B. (1972) A possible precursor of immunoglobulin light chains. Nature New Biol. 239:117-120.

18. Schechter, I. and Burstein, Y. (1976) Partial evidence of the precursors of immunoglobulin light chains of different subgroups: Evidence that the immunoglobulin variable-region gene is larger than hitherto known. Biochem. Biophys. Res. Comm. 68:489-496.

19. Rose, S.M., Kuehl, W.M., and Smith, G.P. (1977) Cloned MPC11 myeloma cells express two kappa genes: a gene for a complete light chain and a gene for a constant region polypeptide. Cell 12:453-462.

20. Pincus, M.R. and Klausner, R.D. (1982) Prediction of the three dimensional structure of the leader sequence of the pre-light chain, a hexadecapeptide. Proc. Natl. Acad. Sci. U.S.A. 79:3413-3417.

21. Hilschmann, N. and Craig, L.C. (1965) Amino acid sequence studies with Bence Jones proteins. Proc. Natl. Acad. Sci. U.S.A. 53:1403-1409.

22. Milstein, C. (1967) Linked groups of residues in immunoglobulin chains. Nature 216:330-332.

23. Kabat, E.A. (1967) Unique features of the variable regions of Bence Jones proteins and their possible relation to antibody complementarity. Proc. Natl. Acad. Sci. U.S.A. 59:613-619.

24. Kabat, E.A. (1970) Heterogeneity and structure of antibody combining sites. Landsteiner Centennial, Dec. 5,6,7, 1968. Ann. N.Y. Acad. Sci. 169:43-54.

25. Franêk, F. (1969) The character of variable sequences in immunoglobulins and its evolutionary origin. In Developmental Aspects of Antibody Formation and Structure. Academia, Czechoslovak Academy of Sciences, Prague, pp. 311-313.

26. Kabat, E.A. (1969) Discussion in Developmental Aspects of Antibody Formation and Structure. Academia, Czechoslovak Academy of Sciences, Prague, pp. 391-393.

27. Kabat, E.A. and Wu, T.T. (1971) Attempts to locate complementarity determining residues in the variable positions of light and heavy chains. Ann. N.Y. Acad. Sci. 190:382-393.

28. Milstein, C. and Pink, J.L.R. (1970) Structure and evolution of immunoglobulins. Prog. Biophys. Mol. Biol. 21:209-263.

29. Capra, J.D., Kehoe, J.M., Winchester, R.J., and Kunkel, H.G. (1971) Structure-function relationships among anti-gamma globulin antibodies. Ann. N.Y. Acad. Sci. 190:371-381.

30. Capra, J.D. and Kehoe, J.M. (1975) Hypervariable regions, idiotypy and the antibody combining site. Advances in Immunology. Academic Press, New York 20:1-40.

31. Thorpe, N.O. and Singer, S.J. (1969) The affinity-labeled residues in antibody active sites. II. Nearest-neighbor analyses. Biochemistry 8:4523-4534.

32. Goetzl, E.J. and Metzger, H. (1970) Affinity labeling of a mouse myeloma protein which binds nitrophenyl ligands. Kinetics of labeling and isolation of a labeled peptide. Biochemistry 9:1267-1278.

33. Franêk, F. (1971) Affinity labeling by m-nitrobenzenediazonium fluoroborate of porcine anti-dinitrophenyl antibodies. Position of labeled tyrosine in the λ-chains. Eur. J. Biochem. 19:176-183.

34. Fleet, G.W.J., Kn
antibody binding site.
photoprecursor of an a

35. Haimovich, J., Eis
Localization of affini
chains of two myelo
Biochemistry 11:2389-2

36. Ray, A. and Cebra,
residues in the primar
raised in strain 13 gu

37. Yoshioka, M., Li
Armstrong, M.Y.K., Kor
Studies on the combir
immunoglobulin which bi
of two types of photo
12:4679-4685.

38. Fisher, C.E. and P
binding site of rabbit
the hypervariable reg
139:135-149.

39. Koo, P.H. and Ceb
distinctive lysyl residu
γ2 chain of guinea pi
Biochemistry 13:184-195

40. Chesebro, B., Hadl
labeling studies on
phosphorylcholine. In
and Cells. 3rd Int. Co
Karger, Basel, pp. 205-

41. Givol, D. (1974) Aff
combining site. Essays

42. Roholt, O.A., Friede
(1973) A light chain ty
rabbit antibenzoate anti

43. Cebra, J.J., Koo,
antibodies: primary stru
186:263-265.

44. Richards, F.F., Li
Konigsberg, W.H. (1974)
region of myeloma prot
labeling patterns. Bioc

45. Padlan, E.A. (1977)
antigen-antibody reacti
diversification of anti
Biophysics 10:35-65.

46. Davies, D.R. and
antibody function. Ann.

47. Kabat, E.A. (1983)
Immunology V, 67-85. Fi
Eds. Yamamura and T. Tad

48. Novotny, J., Bruccole
and Karplus, M. (1983) M
site. J. Biol. Chem. 23:

34. Fleet, G.W.J., Knowles, J.R., and Porter, R.R. (1972) The antibody binding site. Labeling of a specific antibody against the photoprecursor of an aryl nitrène. Biochem. J. 128:499-508.

35. Haimovich, J., Eisen, H.N., Hurwitz, E., and Givol, D. (1972) Localization of affinity-labeled residues on the heavy and light chains of two myeloma proteins with anti-hapten activity. Biochemistry 11:2389-2398.

36. Ray, A. and Cebra, J.J. (1972) Localization of affinity-labeled residues in the primary structure of antidinitrophenyl antibody raised in strain 13 guinea pigs. Biochemistry 11:3647-3656.

37. Yoshioka, M., Lifter, J., Hew, C.-L., Converse, C.A., Armstrong, M.Y.K., Konigsberg, W.H., and Richards, F.F. (1973) Studies on the combining region of protein 460, a mouse γA immunoglobulin which binds several haptens. Binding and reactivity of two types of photoaffinity labeling reagents. Biochemistry 12:4679-4685.

38. Fisher, C.E. and Press, E.M. (1974) Affinity labeling of the binding site of rabbit antibody. Evidence for the involvement of the hypervariable regions of the heavy chain. Biochem J. 139:135-149.

39. Koo, P.H. and Cebra, J.J. (1974) Affinity labeling of a distinctive lysyl residue within the second hypervariable region of γ2 chain of guinea pig anti-p-azobenzenearsonate antibody. Biochemistry 13:184-195.

40. Chesebro, B., Hadler, N., and Metzger, H. (1973) Affinity labeling studies on five mouse myeloma proteins which bind phosphorylcholine. In Specific Receptors of Antibodies, Antigens and Cells. 3rd Int. Convoc. Immunol., 1972, Buffalo, N.Y., S. Karger, Basel, pp. 205-217.

41. Givol, D. (1974) Affinity labeling and topology of the antibody combining site. Essays in Biochemistry 10:1-31.

42. Roholt, O.A., Friedenson, B., Radzimski, G., and Pressman, D. (1973) A light chain tyrosyl sequence in the antibody site of a rabbit antibenzoate antibody. J. Immunol. 111:1367-1375.

43. Cebra, J.J., Koo, P.H., and Ray, A. (1974) Specificity of antibodies: primary structural basis of antibody binding. Science 186:263-265.

44. Richards, F.F., Lifter, J., Hew, C.-L., Yoshioka, M., and Konigsberg, W.H. (1974) Photo- affinity labeling of the combining region of myeloma protein 460. II. An interpretation of the labeling patterns. Biochemistry 13:3572-3574.

45. Padlan, E.A. (1977) Structural basis for the specificity of antigen-antibody reactions and structural mechanisms for the diversification of antigen-binding specificities. Quart. Rev. Biophysics 10:35-65.

46. Davies, D.R. and Metzger, H. (1983) Structural basis of antibody function. Ann. Rev. Immunol. 1:87-117.

47. Kabat, E.A. (1983) The antibody combining site. Progress in Immunology V, 67-85. Fifth International Congress of Immunology. Eds. Yamamura and T. Tada. Academic Press, Tokyo.

48. Novotny, J., Bruccoleri, R.,Newell, J., Murphy, D., Haber, E., and Karplus, M. (1983) Molecular anatomy of the antibody binding site. J. Biol. Chem. 23: 14433-14437.

49. Kabat, E.A., Wu, T.T., and Bilofsky, H. (1977) Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. J. Biol. Chem. 252:6609-6616.

50. Kabat, E.A. (1978) The structural basis of antibody complementarity. Adv. Protein Chem. 32:1-75.

51. Anglister, J. and Zilber, B. (1990) Antibodies against a peptide of cholera toxin differing in cross reactivity with the toxin differ in their specific interactions with the peptide as observed by 1H NMR spectroscopy. Biochemistry 29:921-928.

52. Kabat, E.A., Wu, T.T., and Bilofsky, H. (1978) Variable region genes for the immunoglobulin framework are assembled from small segments of DNA -- A hypothesis. Proc. Natl. Acad. Sci. U.S.A. 75:2429-2433.

53. Hayzer, D.J. and Jaton, J.-C. (1987) Nucleotide Sequence of a cDNA clone encoding a rabbit immunoglobulin-λ light chain: The $V_\lambda$ region differs markedly from that of other species. J. Immunol. 138:2316-2322.

54. Hayzer, D.J. and Jaton, J.-C. (1989) Cloning and sequencing of two functional rabbit germ-line immunoglobulin $V_\lambda$ genes. Gene 80:185-191.

55. Wood, D. and Coleclough, C. (1984) Different joining region J elements of the murine K immunoglobulin light chain locus are used at markedly different frequencies. Proc. Natl. Acad. Sci. U.S.A. 81:4756-4762.

56. Kaartinen, M. and Mäkela, O. (1985) Reading of D genes in variable frames as a source of antibody diversity. Immunology Today 6:324-327.

57. Alt, F.W. and Baltimore, D. (1982) Joining of immunoglobulin heavy chain gene segments: Implications from a chromosome with evidence of three D-$J_H$ fusions. Proc. Natl. Acad. Sci. U.S.A. 79:4118-4122.

58. Perlmutter, R.M., Crews, J.T., Douglas, R., Sorensen, G., Johnson, N., Nivera, N., Gearhart, P.J., and Hood, L. (1984) The generation of diversity in phoshorylcholine-binding antibodies. Adv. in Immunol. Academic Press, New York 35:1-37.

59. Manser, T., Huang, S.Y., and Gefter, M. (1984) Influence of clonal selection on the expression of variable region genes. Science 226:1283-1288.

60. Desiderio, S.V., Yancopoulos, G.D., Paskind, M., Thomas, E., Boss, M.A., Landau, N., Alt, F.W. and Baltimore, D. (1984) Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. Nature 311:752-755.

61. Wysocki, L., Manser, T., Gridley, T. and Gefter, M.L. (1986) Molecular limitations on variable-gene junctional diversity. J. Immunol. 137:3699-3701.

62. Feeney, A.J. (1990) Lack of N regions in fetal and neonatal mouse immunoglobulin V-D-J junctional sequences. J. Exp. Med. 172:1377-1390.

62a. Lafaille, J.J., DeCloux, A., Bonneville, M., Takagaki, Y., and Tonegawa, S. (1989) Junctional sequences of T cell receptor γδ T cell lineage and for a novel intermediate of VDJ joining. Cell 59:859-870.

63. Garcia, I. and Jaton
II and type VII pneumoco
light chains. Immunoc
studies of the light cha

64. Duvoisin, R.M., Ko
Jaton, J.C. (1984) Nuc
constant region of a rab
Eur. J. Immunol. 14:379-

65. Mage, R.G., Bernstei
CHB., Young-Cooper, G.O.
structural and genetic
allotypes of the rabbit.

66. Smith, G.P. (1978)
kappa-chain of mouse mye

67. Edelman, G.M. Cunni
Rutishauser, U., and Wax
an entire γG immunoglobu
63:78-85.

68. Kabat, E.A., Wu, T.
locate residues in compl
combining sites which m
Proc. Natl. Acad. Sci. U

69. Padlan, E.A., Davies
Chem. (1976) Model buildin
hapten-binding site of M
Biol. 41:627-637.

70. Kabat, E.A. (1982)
complementarity. Pharma

71. Dwek, R.A., Wain-Hob
B., Perkins, S.J., and G
combining site by magnet

72. Feldman, R.J., Pott
hypothetical space-filli
binding myeloma J539. M

73. Givol, D. (1979) The
Recognition II (E. Lenno
11B, 23:71-125.

74. Padlan, E.A. and Ka
combining sites. in Mol
and Applications. Meth
Abelson and M.I. Simon,
Cyclo

75. Kabat, E.A. and Wu,
and segments of sequences
Relative contributions o
binding of antibody comb
Acad

76. Gibbs, M.R., Moody,
structure of the asp
chloramphenicol acetyltra
consequences of disruptio
29:11261-11265.
Immun

77. Kriegler, M. (1990)
Press, New York, pp.1-24(
Cloning

78. Horwitz, A.H., Chang,
Secretion of functional a
Proc. Natl. Acad. Sci. US

63. Garcia, I. and Jaton, J.C. (1979) The Immune response to type II and type VII pneumococcal vaccines in Basilea rabbits lacking K light chains. Immunochemical and partial amino acid sequence studies of the light chains. Mol. Immunol. 16:1063-1071.

64. Duvoisin, R.M., Kocher, H.P., Garcia, I., Rougeon, F. and Jaton, J.C. (1984) Nucleotide sequence of a cDNA encoding the constant region of a rabbit immunoglobulin light chain of λ type. Eur. J. Immunol. 14:379-382.

65. Mage, R.G., Bernstein, K.E., McCartney-Francis, N., Alexander, C.B., Young-Cooper, G.O., Padlan, E.A., and Cohen, G.H. (1984) The structural and genetic basis for expression of normal and latent allotypes of the rabbit. Mol. Immunol. 21:1067-1081.

66. Smith, G.P. (1978) Sequence of the full-length immunoglobulin kappa-chain of mouse myeloma MPC 11. Biochem. J. 171:337-347.

67. Edelman, G.M. Cunningham, B.A. Gall, W.E., Gottlieb, P.D., Rutishauser, U., and Waxdal, M.J. (1969) The covalent structure of an entire γG immunoglobulin molecule. Proc. Natl. Acad. Sci. USA 63:78-85.

68. Kabat, E.A., Wu, T.T., and Bilofsky, H. (1976) Attempts to locate residues in complementarity determining regions of antibody combining sites which make contact with antigenic determinants. Proc. Natl. Acad. Sci. U.S.A. 73:617-619.

69. Padlan, E.A., Davies, D.R., Pecht, I., Givol. D., and Wright, C. (1976) Model building studies of antigen-binding sites: the hapten-binding site of MOPC 315. Cold Spring Harbor Symp. Quant. Biol. 41:627-637.

70. Kabat, E.A. (1982) Antibody diversity versus antibody complementarity. Pharmacol. Reviews 34:23-38.

71. Dwek, R.A., Wain-Hobson, S., Dower, S., Gettins, P., Sutton, B., Perkins, S.J., and Givol, D. (1977) Structure of an antibody combining site by magnetic resonance. Nature 266:31-37.

72. Feldman, R.J., Potter, M., and Glaudemans, C.P.J. (1981) A hypothetical space-filling model of the V-regions of the galactan binding myeloma J539. Mol. Immunol. 18:683-698.

73. Givol, D. (1979) The antibody combining site. In Defense and Recognition II (E. Lennox, Ed.) Univ. Park Press, Baltimore, MD, 11B, 23:71-125.

74. Padlan, E.A. and Kabat, E.A. (1991) Modelling of antibody combining sites. in Molecular Design and Modelling. in Concepts and Applications. Methods in Enzymology Academic Press. J.N. Abelson and M.I. Simon, Editors in chief.

75. Kabat, E.A. and Wu, T.T. (1991) Identical V-region sequences and segments of sequences in antibodies of different specificities-Relative contributions of $V_H$ and $V_L$ genes, minigenes and CDRs to binding of antibody combining sites. J. Immunol. (in press).

76. Gibbs, M.R., Moody, P.C.E., and Leslie, A.G.W. (1990) Crystal structure of the aspartic acid-199→asparagine mutant of chloramphenicol acetyltransferase to 2.35-A resolution: structural consequences of disruption of a buried salt bridge. Biochemistry 29:11261-11265.

77. Kriegler, M. (1990) Gene Transfer and Expression. Stockton Press, New York, pp.1-240.

78. Horwitz, A.H., Chang, C.P., Better, M. and Hellstrom, K. (1988) Secretion of functional antibody and Fab fragment from yeast cells. Proc. Natl. Acad. Sci. USA 8678-8682.